

????

???

□ (□) □□	□□
???? (Artificial Intelligence)	Marvin Minsky: AI ?????????????????????? ???AI?: AI ??????????????????
???? (Generalization Ability)	????????????????????????????
??? (Hyperparameter)	?? ?????????????
????? (Transfer Learning)	???????? AI ?? ????????????????????????????????
?? (Hallucination)	???????????????????? (?: ??????????) ?? ????????????????????
????? (Model Distillation)	?? ????????????????????????????
TTQA (????????????)	?? AI ???

??? vs ??

- ?????????????????????????????
- ?????????????????????????????

??

?????????????? $f(x) = ax + b = y$

- ?? x????????????????????
- ?? a, b????????
- ?? y????????????????...

??? vs ??? vs ???

- ?????????AI ??
- ??????AI ???
- ?????????AI ??????????????????

??

2. **???** (Bar Chart)????????????????????

3. **????Scatter plot**??x????????y
??

??

3. **????Radar chart?**

??
??????

??

4. **???** (Box chart)????????????????????????????????

5. **???** (Line Chart)????????????????????????????

6. **???** (Pi Chart)???????????????? 100% ??????

????

?????Quartile??25%????????????

- ???????Q1??????25%??
- ???????Q2??????50%??
- ???????Q3??????25%??
- ?????IQR??Q3 - Q1????????

??Q1 = 60?Q2 = 70?Q3 = 80?IQR = 20
????????????70??25%????60????25%????80????????????20?

???? (IQR) ???

- ????????
- ????????
- ??? 50% ????????

?????

??? Q1?Q2 ? Q3????????????

1. ?????????????80, 85, 85, 90, 90, 95, 95, 100, 100

2. ?????????

- ?????? (Q1) ????????? 25% ??????
- ?????? (Q2) ????????? 50% ??????????
- ?????? (Q3) ????????? 75% ??????

3. ?????? 9 ?????????

- 25% ?? = (9 + 1) * 0.25 = 2.5????? 3??? Q1 ?? 3 ???
- 50% ?? = (9 + 1) * 0.5 = 5??? Q2 ?? 5 ???
- 75% ?? = (9 + 1) * 0.75 = 7.5???? Q3 ?? 7 ???

4. ?????????????????

- Q1 = ? 3 ?? = 85
- Q2 = ? 5 ?? = 90
- Q3 = ? 8 ?? = 100

???Q1 = 85?Q2 = 90?Q3 = 95?

???? (IQR) = Q3 - Q1 = 95 - 85 = 10

DALL-E vs Midjourney

DALL-E ? Midjourney ?????????????????????????????????

DALL-E

- DALL-E ?????????????VAE??_transformer ?????????
- ?????????????????????????????????
- DALL-E ?????????????????????????????????
- DALL-E ?????????????????????????????????

Midjourney

- Midjourney ?????????????GAN??_diffusion ?????????
- ?????????????????????????????????
- Midjourney ????????????????????? DALL-E ?????????
- Midjourney ?????????????????????????????????

??

- ?????DALL-E ? Midjourney ????????????? DALL-E ?????????????
- ?????Midjourney ? DALL-E ?????????????
- ?????DALL-E ? Midjourney ????????????? DALL-E ?????????????
- ?????DALL-E ?????????????????????????????????Midjourney
????????????????????????????????

???DALL-E ? Midjourney ?????????????????????????????

BERT (Transformer)

BERT (Transformer) ??????

1. ??Transformer??BERT ???Transformer????????????Self-Attention
???????????? BERT ?????????????????????????
2. ???BERT ?????????????????????????????????????Masked
Language Modeling????????Next Sentence Prediction??
3. ?????????????BERT ?????????????????????????????????

?????

1. ?????BERT ?????????????????????????
2. ?????BERT ?????????????????????????
3. ?????BERT ?????????????????????????
4. ?????BERT ?????????????????????????

BERT ????????????????????????????????????? NLP ??????????

GPT vs BERT

GPT is one directional while BERT is bidirectional. You use GPT to predict the next token GIVEN the previous context, while you use BERT to predict the token in between GIVEN the token of both previous and future context. So you use GPT for **text generation** while you use BERT for **text classification** (understanding text). In terms of architecture, GPT is **transformer-decoder** only while BERT is **transformer-encoder** only.

???/???/???/???

?????

- FP(??) / TP(??) / ??(FN) / ??(TN)
- Accuracy????????????????????????
 - $Accuracy = (TP+TN) / (TP+FP+FN+TN)$
- Precision????????????????????FP????????
 - $Precision = TP / (TP+FP)$
- Recall-TPR????????????????????FN????????
 - $Recall = TP / (TP+FN)$
- Error????????????
 - $Error = (FP+FN) / (TP+FP+FN+TN)$
 - $Error = 1 - Accuracy$
- FPR(FPR)FP / (FP+TN)
- F1 Score?Precision ? Recall ??????????????????????
- ROC ?????????????????????(????)????????????(????-????)
????????????????????????????????
- AUC??ROC????????????0.5????????????????????????AUC?????1
????????????????????AUC??0.7????????????0.8????????0.9????????

- **???** MSE/MAE Outliers
4. **???** (R-Square)

?? (Learning Rate)

?? (Learning Rate)

???

- **???** ?
- **???** ?

???

1. 10 ~ 100
2. **Adam / AdamW** $1e^{-3}$
3. 1/10–1/100

???

- **???**
- **???**
- **??**
- **???**

LLM (Benchmark)

- [\[Day18\] LLM Benchmark ? Leaderboard - iT](#)

???	??	?????	??????
MMLU	???	?? 57	????????????
HellaSwag	???	????????????	????????????
ARC	???	????????????	????????????
HumanEval	???	??????????	????????????
GSM8K	???	????????????	????????????

Machine Learning Training Data Overfitting

????????????

1. Fold??
2. Test Data????????
3. ?????????
4. 2-3????????
5. ?????????

??? 120 ??? K = 5?5 ???

1. ??
 - ? 120 ??? 5 ??? 24 ??
 - 1: 1?24?2: 25?48?3: 49?72?4: 73?96?5: 97?120?
2. ??
 - ????
 - ??? 2?5 ? (25?120)
 - ??? 1 ? (1?24)
 - ????
 - ??? 1?3?5 ?
 - ??? 2 ?
 - ... ????? 5 ?????
3. ??
 - ????????? MSE?Accuracy??
 - ??? 5 ????????? CV ???
 - ??? ??? ?????????

????????

1. ?????????????????
2. ?????????????????
3. ?????????????????

????????

- K?????K-Fold Cross-Validation?????K????????
- ??????Holdout Method????????

??? (Batch Inference)

???????? (?????) ??????

GRPO ???

- ??????? OpenAI o1 ????????
- ????????
- ????????

GDPR

?????????(GDPR) ? AI ????

- ????????
- ????????
- ????????
- ?? AI ????????
- ??? AI ??????

??? (Overfitting)

??? (Overfitting)	??? (Underfitting)
???????????	????
<ul style="list-style-type: none">• ??? (L1/L2)• Dropout (???????)• Early Stopping (???)• Data Augmentation (????)	<ul style="list-style-type: none">• ??????• ????• ????????• ????????

??? vs ???

???????????

????

- ??????? 0 ~ 1 ??
- ??????????(? cm ? kg)????????????????????(KNN)

????

- ??????????? 0?????? 1 ???